Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Linear Algebra and its Applications 410 (2005) 279–287
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Matrix equations with restraints and their statistical applications

Czesław Stępniak *

Institute of Mathematics, University of Rzeszów, Al. Rejtana 16 A, PL-35-959 Rzeszów, Poland
Department of Statistics and Econometrics, Faculty of Economics, Maria Curie-Skłodowska University, Pl. Marii Curie-Skłodowskiej 5, PL-20-031 Lublin, Poland

Received 22 August 2002; accepted 26 August 2002

Submitted by H.J. Werner

Dedicated to Professor George Styan for his 65th birthday

Abstract

Matrix equations with various conditions suited for statistical purposes in linear experiments are considered.

Keywords: Matrix equation; Restraints; Linear normal model; Quadratic plus linear estimation; Comparison of models

1. Notation and statistical background

The classical vector–matrix notation is used here. Among others, if M is a matrix, then M', $R(M)$, $N(M)$, $r(M)$ and P_M denote, respectively, the transposition, the range (column space), the kernel (null space), the rank of M, and the orthogonal projector on $R(M)$. By M^-, M^+ and $M^\perp$, respectively, is denoted a generalized inverse,

* Address: Institute of Mathematics, University of Rzeszów, Al. Rejtana 16 A, PL-35-959 Rzeszów, Poland.
E-mail address: cees@univ.rzeszow.pl

0024-3795/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2002.08.001

the Moore–Penrose generalized inverse of M and a matrix of maximum rank such that $MM^{\perp} = 0$. The symbol $M \geqslant 0$ means that M is nonnegative definite (n.n.d.). Moreover, the symbol R^n stands for the n-dimensional Euclidean space represented by column vectors.

Let us consider a random vector x in R^n. Suppose x is subject to a normal linear model $N(A\beta, \sigma V)$, where A and V are known matrices, while $\beta \in R^p$ and $\sigma > 0$ are unknown parameters. We shall say that the model has trivial deterministic part if for any $a \in R^n$, $\mathrm{var}(a'x) = 0$ implies $E(a'x) = 0$. Such a model is also known in the literature as a weakly singular model. It is known that $N(A\beta, \sigma V)$ has trivial deterministic part, if and only if, $R(A) \subseteq R(V)$ (cf. [11, Lemma 3]).

Assume $N(A\beta, \sigma V)$ has trivial deterministic part. Then there exists a nonsingular linear transformation $z = Fx$ and a reparametrization $\theta = C\beta$ such that z is subject to the model

$$N\left(\begin{bmatrix} \theta \\ 0_{n-r} \end{bmatrix}, \sigma \begin{bmatrix} I_q & 0 \\ 0 & 0_{n-q} \end{bmatrix}\right),$$

where $r = r(A)$ and $q = r(V)$.

Denote by z_i, $i = 1, \dots, n$, the components of the random vector z and assume that $r(A) < r(V)$. Then the vector statistic $(z_1, \dots, z_r)$ and the scalar statistic $s = \frac{1}{q-r}\sum_{i=r+1}^{q} z_i^2$ are the *Best Unbiased Estimators* of θ and σ, respectively (see [11, Theorem 6]). Moreover, by Lehmann [3, p. 142] the statistics $(z_1, \dots, z_r)$ and s are complete and sufficient.

Now let us consider two random vectors $x \in R^n$ and $y \in R^m$. Suppose x is subject to a normal linear model $N(A\beta, \sigma V)$ and y is subject to a normal linear model $N(B\beta, \sigma W)$. In this context model $N(A\beta, \sigma V)$ is said to be *at least as good as* the model $N(B\beta, \sigma W)$, if for any parametric function $\psi = k'\beta + c\sigma$ and for any unbiased estimator $\hat{\psi} = \hat{\psi}(y)$, whenever such exists, there exists an unbiased estimator $\tilde{\psi} = \tilde{\psi}(x)$ such that $\mathrm{var}(\tilde{\psi}) \leqslant \mathrm{var}(\hat{\psi})$ for all β and σ.

Since the minimal sufficient statistics in the models $N_1 = N(A\beta, \sigma V)$ and $N_2 = N(B\beta, \sigma V)$ are complete, the first model is at least as good as the second one, if and only if,

- N_1 is at least as good as N_2 with respect to linear estimation, and,
- The number of degrees of freedom for error in N_1 is not less than in N_2.

Thus, by Stępniak [9,10] the necessary and sufficient conditions for N_1 to be at least as good as N_2 can be written in the following algebraic form:

$$\text{There exists a linear transform } F \text{ such that } B = FA \text{ and } W - FVF' \text{ is n.n.d.,} \tag{1}$$

and

$$r(V + AA') - r(A) \geqslant r(W + BB') - r(B). \tag{2}$$

The problem of characterization of these conditions in terms of linear transformation of the observation vector is the groundwork for this paper. This problem was stimulated by earlier works by Baksalary and Kala [1], Drygas [2], Mueller [5], Oktaba et al. [7], Stępniak [10,12] and Torgersen [13]. However, as yet, the known results in the subject are not satisfactory from algebraic point of view. In particular, they do not explain the nature of the assumptions that were undertaken.

Not only the present paper throws a new light onto this problem, but it also contributes to the matrix algebra, as well. Moreover some efforts are taken to make this paper as self-contained as possible.

2. Auxiliary lemmas

Lemma 1. *Let M_1 and M_2 be arbitrary matrices with the same number of rows, and let P_1 and P_2 be the orthogonal projectors on $R(M_1)$ and $R(M_2)$, respectively. Then*

$$r(M_1'M_2) = r(P_1P_2).$$

Proof. Really,

$$\begin{aligned} r(M_1'M_2) &= \dim R(M_1'M_2) = \dim R(M_1'P_2) \\ &= r(M_1'P_2) = r(P_2M_1) = \dim R(P_2P_1) \\ &= r(P_1P_2). \qquad \square \end{aligned}$$

Lemma 2. *Let A, V and F be arbitrary matrices with the same number of rows, such that $R(A) \subseteq R(V)$. Then*

$$r(V) - r(A) \geqslant r(F'V) - r(F'A).$$

Proof. Denote by P_M the orthogonal projector onto $R(M)$. Then, by Lemma 1,

$$\begin{aligned} r(F'V) - r(F'A) &= r(P_FP_V) - r(P_FP_A) \\ &= \dim R(P_FP_V) - \dim R(P_FP_A) \\ &\leqslant \dim R[P_F(P_V - P_A)] \\ &= r[P_F(P_V - P_A)] \\ &\leqslant r(P_V - P_A). \end{aligned}$$

On the other hand $P_V = (P_V - P_A) - P_A$ and, via assumption $R(A) \subseteq R(V)$, $(P_V - P_A)P_A = 0$. Thus $r(P_V - P_A) = r(V) - r(A)$, completing the proof. $\square$

Remark 1. Lemma 2 may also be proved in an another way, by using Marsaglia and Styan [6, Corollary 6.2].

Lemma 3. *Let M_1 and M_2 be symmetric nonnegative definite matrices of the same order.*

(a) *If $M_2 - M_1$ is n.n.d. and $r(M_1) \geqslant r(M_2)$ then $R(M_1) = R(M_2)$.*
(b) *If $R(M_1) \subseteq R(M_2)$ then there exists a positive scalar c such that $M_2 - cM_1$ is n.n.d.*

Proof. (a) It follows by the evident fact that $M_2 - M_1 \geqslant 0$ implies $R(M_1) \subseteq R(M_2)$.

(b) We only need to use the fact that $M_2 - M_1 \geqslant 0$, if and only if, $R(M_1) \subseteq R(M_2)$ and the maximal eigenvalue of $M_2^+ M_1$ is not greater than 1 (cf. [9, Theorem 1]). □

3. Matrix equations with restraints

For given matrices A and B with the same number of columns consider the matrix equation

$$B = XA. \tag{3}$$

It is well known that this equation is consistent, if and only if, $R(B') \subseteq R(A')$ and, if so, then its solution is usually not unique. In this situation some restrains on the solution may be posed. In particular, we are interested in the restrain $R(VX') \subseteq R(A)$, where V is a symmetric n.n.d. matrix.

Theorem 4. *Let A and B be arbitrary matrices with the same number of columns, and let V be a symmetric n.n.d. matrix such that $R(A) \subseteq R(V)$.*

(a) *If $B = FA$ for some F then*

$$F_0 = FV^{\frac{1}{2}} P_{(V^+)^{\frac{1}{2}}A} (V^+)^{\frac{1}{2}} \tag{4}$$

satisfies the conditions

$$B = F_0 A \quad \text{and} \quad R(VF_0') \subseteq R(A). \tag{5}$$

(b) *If the condition (5) holds then for any matrix F_1 of the same dimension as F_0, $B = F_1 A$ and $R(VF_1') \subseteq R(A)$ if and only if $R(F_1' - F_0') \subseteq N(V)$.*
(c) *If the condition (5) holds then $R(F_0 V) = R(B)$ and $FVF' - F_0 V F_0'$ is n.n.d. for arbitrary F satisfying $B = FA$.*

Proof. (a) Since $V^{\frac{1}{2}}(V^+)^{\frac{1}{2}} = P_V$ and $R(A) \subseteq R(V)$, we have the identity $V^{\frac{1}{2}} P_{(V^+)^{\frac{1}{2}}A} (V^+)^{\frac{1}{2}} A = A$. Let $B = FA$ and F_0 be defined by (4). Then $F_0 A = FA = B$ and $R(VF_0') \subseteq R\big[V(V^+)^{\frac{1}{2}} P_{(V^+)^{\frac{1}{2}}A} V^{\frac{1}{2}}\big] \subseteq R\big[V(V^+)^{\frac{1}{2}}(V^+)^{\frac{1}{2}} A\big] = R(A)$.

(b) Suppose the conditions (5), $B = F_1A$ and $R(VF_1') \subseteq R(A)$ hold. Then $(F_1 - F_0)A = 0$ and $R[V(F_1' - F_0')] \subseteq R(A)$. Writing the first one in the form $R(F_1' - F_0') \subseteq N(A')$ and using the second one, we get $(F_1 - F_0)V(F_1' - F_0') = 0$. The last one can be rewritten in the form $V(F_1' - F_0') = 0$, implying $R(F_1' - F_0') \subseteq N(V)$.

Conversely, if $R(F_1' - F_0') \subseteq N(V)$ then $V(F_1' - F_0') = 0$, and, in particular, $(F_1 - F_0)A = 0$. Therefore, by (5), F_1 satisfies the desired conditions $B = F_1A$ and $R(VF_1') \subseteq R(A)$.

(c) Since $R(B) = R(F_0A)$ and $R(A) \subseteq R(V)$, the relation $R(B) \subseteq R(F_0V)$ is evident. Thus it remains to verify that $R(F_0V) \subseteq R(F_0A)$ or, equivalently, that

$$N(A'F_0') \subseteq N(VF_0'). \tag{6}$$

Really, if $x \in N(A'F_0')$ then $F_0'x \in N(A')$ and, by $R(VF_0') \subseteq R(A)$, $F_0'x \in N(F_0V)$. In consequence $F_0VF_0'x = 0$, and hence $x \in N(VF_0')$, completing the proof of the inclusion (6) and the proof of the equality $R(F_0V) = R(B)$.

In order to prove that $FVF' - F_0VF_0'$ is n.n.d. let us use the evident fact that $MM' - MPM'$ is n.n.d. for any matrix M and for any orthogonal projector P. Now we only need to set $M = FV^{\frac{1}{2}}$ and $P = P_{(V^+)^{\frac{1}{2}}A}$. □

Now let us consider the matrix equation $B = XA$ under restrain $W - XVX' \geqslant 0$, where V and W are symmetric n.n.d. matrices. If the conditions

$$B = XA \quad \text{and} \quad W - XVX' \text{ is n.n.d.}, \tag{7}$$

are consistent, then a solution of (7) is usually not unique. For some statistical reasons we are interested in a solution of (7) satisfying the condition $R(XV) = R(W)$.

Theorem 5. *If the conditions* (7) *are consistent and, moreover,* $r(V) - r(A) \geqslant r(W) - r(B)$, $R(A) \subseteq R(V)$ *and* $R(B) \subseteq R(W)$, *then there exists a matrix* F *such that*

$$B = FA, \quad W - FVF' \textit{ is n.n.d.} \quad \textit{and} \quad R(FV) = R(W). \tag{8}$$

Proof. By the assumption $R(A) \subseteq R(V)$ the orthogonal projector onto $R(A)^\perp \cap R(V)$ has a sense and it may be presented in the form $P_V - P_A$. Similarly, by the assumption $R(B) \subseteq R(W)$ the orthogonal projector onto $R(B)^\perp \cap R(W)$ may be presented in the form $P_W - P_B$.

Let $c_1, \ldots, c_r$, where $r = r(W) - r(B)$, be an orthonormal basis in $R(P_W - P_B)$, and $d_1, \ldots, d_r$ be an orthonormal basis in an r-dimensional subspace of $R(P_V - P_A)$. Such a possibility follows by the assumption $r(V) - r(A) \geqslant r(W) - r(B)$. Introduce an $m \times n$ matrix

$$F_1 = CD', \tag{9}$$

where $C = [c_1, \ldots, c_r]$ and $D = [d_1, \ldots, d_r]$. We note that

$$F_1A = 0. \tag{10}$$

Suppose (7) is consistent. Then, by Theorem 4(c), there exists a matrix F_0 such that $B = F_0A$, $R(VF_0') \subseteq R(A)$, $W - F_0VF_0' \geqslant 0$ and

$$R(F_0V) = R(B). \tag{11}$$

The inclusion $R(VF_0') \subseteq R(A)$ implies

$$F_1VF_0' = 0. \tag{12}$$

On the other hand, by conditions $W - F_0VF_0' \geqslant 0$ and $R(F_0V) = R(B)$, we get

$$R(W - F_0VF_0') \supseteq R(F_1) \supseteq R(F_1VF_1').$$

Thus, by Lemma 3(b), there exists a scalar $c > 0$ such that

$$W - F_0VF_0' - cF_1VF_1' \geqslant 0. \tag{13}$$

Let us set

$$F = F_0 + \sqrt{c}F_1. \tag{14}$$

The condition $B = FA$ follows by the facts $F_0A = B$ and $F_1A = 0$. Moreover, via (12),

$$FVF' = F_0VF_0' + cF_1VF_1', \tag{15}$$

and hence, by (13) and (15), we get $W - FVF' \geqslant 0$. Thus it remains to verify that $R(FV) = R(W)$.

At first we shall show that $r(FVF') \geqslant r(W)$. Really, by (12),

$$r(FVF') = r(F_0VF_0') + r(F_1VF_1').$$

It follows from (11) that $r(F_0VF_0') = r(B)$. On the other hand,

$$\begin{aligned} r(F_1VF_1') &\geqslant r(F_1'F_1VF_1'F_1) = r(DD'VDD') \\ &= r(P_DV) = r(P_DP_V) = r(P_D) \\ &= r(W) - r(B). \end{aligned}$$

Thus $r(FVF') \geqslant R(W)$, and the desired condition $R(FV) = R(W)$ follows from (13) and (15), via Lemma 3(a).

In this way Theorem 5 is proved. □

Now we are ready to present the main result in this section.

Let A and B be arbitrary matrices of dimension $n \times p$ and $m \times p$, and let V and W be symmetric n.n.d. matrices of order n and m, respectively. Consider the following conditions:

(i) There exists an $m \times n$ matrix F such that $B = FA$, $W - FVF'$ is n.n.d. and $r(V) - r(A) \geqslant r(W) - r(B)$.

(ii) There exists an $m \times n$ matrix F such that $B = FA$ and a symmetric n.n.d. matrix G of order m such that $R(G) \subseteq R(FV)$ and

$$FVF' + G = W.$$

(iii) There exists an $m \times n$ matrix F such that $B = FA$ and a symmetric n.n.d. matrix H of order n such that $R(H) \subseteq R(V)$ and

$$F(V + H)F' = W.$$

Theorem 6. *Under the above assumptions:*

(a) *The conditions* (ii) *and* (iii) *are equivalent.*
(b) *If $R(A) \subseteq R(V)$ then each of the conditions* (ii) *or* (iii) *implies* (i).
(c) *If $R(A) \subseteq R(V)$ and $R(B) \subseteq R(W)$ then the conditions* (i)–(iii) *are equivalent.*

Proof. (a) (iii) $\Rightarrow$ (ii). After setting $G = FHF'$ we only need to verify that $R(G) \subseteq R(FV)$. The last follows immediately from the facts that $R(G) \subseteq R(FH)$ and $R(H) \subseteq R(V)$.

(ii) $\Rightarrow$ (iii). Suppose (ii) holds. We only need to show that G can be presented in the form FHF' for some H satisfying $R(H) \subseteq R(V)$. Really, if $R(G) \subseteq R(FV)$, then $R(G^{\frac{1}{2}}) \subseteq R(FV)$. Thus there exists a matrix C such that $G^{\frac{1}{2}} = FVC$ and, in consequence, $G = FHF'$ with $H = VCC'V$. In this way (a) is proved.

(b) We observe that (ii) implies $B = FA$, $R(W) = R(FV)$ and $W - FVF'$ is n.n.d. Thus we only need to show that $r(V) - r(A) \geqslant r(W) - r(B)$. By Lemma 2, via assumption $R(A) \subseteq R(V)$, we get

$$r(V) - r(A) \geqslant r(FV) - r(FA) = r(W) - r(B),$$

completing the proof of (b).

(c) It follows directly by Theorem 5 via (a) and (b). □

4. Applications in linear experiments

Let x be observation vector in a linear model with expectation $\mu = A\beta$ and the variance–covariance matrix σV. Then by the well known Lehmann–Scheffé Theorem (see [4] or [8]) a vector statistic Fx is a *Best Unbiased Estimator* of μ, if and only if, it is not correlated with any unbiased estimator of zero. This leads to the conditions: $FA = A$ and $R(VF') \subseteq R(A)$ (cf. [14]). The following theorem will be convenient in the further consideration.

Theorem 7. *Let x be a random vector with expectation $\mu = A\beta$ and the variance–covariance matrix σV. If $R(A) \subseteq R(V)$ then*

(a) *The statistic*

$$\hat{\mu} = V^{\frac{1}{2}} P_{(V^+)^{\frac{1}{2}} A} (V^+)^{\frac{1}{2}} x$$

is a Best Unbiased Estimator of μ.
(b) *If $\tilde{\mu}$ is any Best Unbiased Estimator of μ then $\tilde{\mu} = \hat{\mu}$ almost surely.*

Proof. Assertion (a) follows from Theorem 4(a) by setting $B = A$ and $F = I$, while the assertion (b) follows directly from Theorem 4 (b). □

Now consider normal linear models $N_1 = N(A\beta, \sigma V)$ and $N_2 = N(B\beta, \sigma W)$. Let us remind that necessary and sufficient conditions for N_1 to be at least as good as N_2 (for estimation of all functions $\psi = k'\beta + c\sigma$) reduce to (1) and (2). The following theorem characterizes these conditions in terms of linear transformations of the observation vector.

Theorem 8. *Let $x \in R^n$ and $y \in R^m$ be observation vectors in normal linear models $N_1 = N(A\beta, \sigma V)$ and $N_2 = N(B\beta, \sigma W)$, respectively. Then the following conditions are equivalent*:

(i) *The model N_1 is at least as good as the model N_2.*
(ii) *There exists an $m \times n$ matrix F and a normal random vector $z \in R^m$, independent of x, with expectation zero and the variance–covariance matrix σG, such that $R(G) \subseteq R[F(V + AA')]$ and $Fx + z$ has the same distribution as y.*
(iii) *There exists an $m \times n$ matrix F and a normal random vector $u \in R^n$, independent of x, with expectation zero and the variance–covariance matrix σH, such that $R(H) \subseteq R(V + AA')$ and $F(x + u)$ has the same distribution as y.*

Proof. Let us define $V_1 = V + AA'$ and $W_1 = W + BB'$. We observe that the conditions $W - FVF' \geqslant 0$ and $W_1 - FV_1F' \geqslant 0$ are equivalent providing that $B = FA$, and, moreover, $R(A) \subseteq R(V_1)$ and $R(B) \subseteq R(W_1)$. Now the Theorem follows directly from Theorem 6(c). □

Remark 2. If $R(A) \subseteq R(V)$ then the term $V + AA'$ in (ii) and (iii) may be replaced by V.

5. Acknowledgment

The author wishes to thank the referee for his careful reading and valuable comments.

References

[1] J.K. Baksalary, R. Kala, Linear transformations preserving best linear unbiased estimators in a general Gauss–Markoff model, Ann. Statist. 9 (1981) 913–916.
[2] H. Drygas, Sufficiency and completeness in the Gauss–Markov model, Sankhyā Ser. A 45 (1983) 88–98.
[3] E.L. Lehmann, Testing Statistical Hypotheses, second ed., Wiley, New York, 1986.
[4] E.L. Lehmann, H. Scheffé, Completeness, similar regions and unbiased estimation, Sankhyā 10 (1950) 305–340.

[5] J. Mueller, Sufficiency and completeness in the linear model, J. Multivariate Anal. 21 (1987) 312–323.
[6] G. Marsaglia, G.P.H. Styan, Equalities and inequalities for ranks of matrices, Linear and Multilinear Algebra 2 (1974) 269–292.
[7] W. Oktaba, A. Kornacki, J. Wawrzosek, Invariant linearly sufficient transformations of the general Gauss–Markoff model. Estimation and testing, Scand. J. Statist. 15 (1988) 117–124.
[8] S. Puntanen, G.P.H. Styan, H.J. Werner, Two matrix-based proofs that the linear estimator Gy is the best linear unbiased estimator, J. Stat. Plann. Inference 88 (2000) 173–179.
[9] C. Stępniak, Ordering of nonnegative definite matrices with application to comparison of linear models, Linear Algebra Appl. 70 (1985) 67–71.
[10] C. Stępniak, Comparing normal linear experiments and transformations of observations, Statistics 30 (1998) 279–289.
[11] C. Stępniak, Geometry of linear and quadratic estimation in a normal linear model with singular variance–covariance matrix, Demonstratio Math. 32 (1999) 639–645.
[12] C. Stępniak, On a matrix result in comparison of linear experiments, Linear Algebra Appl. 321 (2000) 321–325.
[13] E. Torgersen, Comparison of Statistical Experiments, Cambridge Univ. Press, Cambridge, 1991.
[14] G. Zyskind, On canonical forms, non-negative covariance matrices and best and simple least squares estimators in linear models, Ann. Math. Statist. 38 (1967) 1092–1109.